

# Multi-scale spatial modeling of human exposure from local sources to global intake

**Wannaz, Cedric; Fantke, Peter; Jolliet, Olivier**

*Published in:*
Environmental Science and Technology

*Link to article, DOI:*
10.1021/acs.est.7b05099

*Publication date:*
2018

*Document Version*
Peer reviewed version

# Multi-scale spatial modeling of human exposure from local sources to global intake

Cedric Wannaz[a,*], Peter Fantke[b], Olivier Jolliet[a]

[a] School of Public Health (SPH), University of Michigan, 6622 SPH Tower, 1415 Washington Heights, Ann Arbor, Michigan 48109-2029, United States

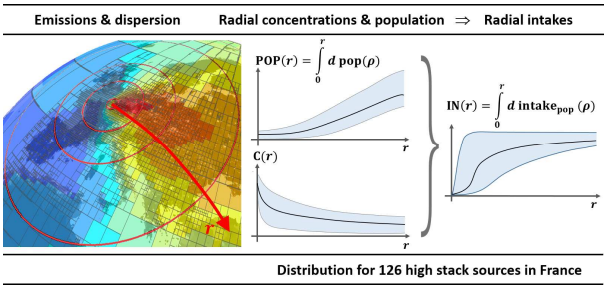
[b] Quantitative Sustainability Assessment Division, Department of Management Engineering, Technical University of Denmark, Bygningstorvet 116, 2800 Kgs. Lyngby, Denmark

[*]Corresponding author: phone: (+1)734-548-2535, E-mail address: wannaz@umich.edu.

# Abstract

Exposure studies, used in human health risk and impact assessments of chemicals are largely performed locally or regionally. It is usually not known how global impacts resulting from exposure to point source emissions compare to local impacts. To address this problem, we introduce *Pangea*, an innovative multi-scale, spatial multimedia fate and exposure assessment model. We study local to global population exposure associated with emissions from 126 point sources matching locations of waste-to-energy plants across France. Results for three chemicals with distinct physicochemical properties are expressed as the evolution of the population intake fraction through inhalation and ingestion as a function of the distance from sources. For substances with atmospheric half-lives longer than a week, less than 20% of the global population intake through inhalation (median of 126 emission scenarios) can occur within a 100 km radius from the source. This suggests that, by neglecting distant low-level exposure, local assessments might only account for fractions of global cumulative intakes. We also study ~10,000 emission locations covering France more densely to determine per chemical and exposure route which locations minimize global intakes. Maps of global intake fractions associated with each emission location show clear patterns associated with population and agriculture production densities.


**Keywords:** spatial multimedia model; multi-pathway exposure assessment; population intake fraction; environmental fate of chemicals; local to global modeling




[Table of Contents (TOC)/Abstract Art]

# INTRODUCTION

For human health risk and impact assessments of chemical emissions, exposure studies either focus on local or regional scales, or assess larger areas at the expense of spatial resolution. How impacts on human health at the global scale compare to local impacts for the same emission source has not been systematically studied. We therefore focus on assessing the spatial distribution of the *intake fraction*, the fraction of an emission that is ultimately taken in by an exposed human population, from local to global scale.

Exposure studies in risk or health impact assessments generally focus on regions with high pollutant concentrations and model fate and transport of pollutants within close proximity of emission sources, typically within a radius of 5 km to 30 km.[1–3] Such studies neglect potentially important intakes associated with populations exposed on a larger geographical scale. Lohman and Seigneur[4] show for example that up to 90% of polychlorinated dibenzo-*p*-dioxins and polychlorinated dibenzofurans (PCDD/F) emitted from stacks might be deposited farther than 100 km away from sources. Other studies analyzing long range transport[5–12] also suggest that local assessments might only account for a fraction of global exposure, especially for chemicals that are both mobile and persistent.

In contrast, exposure studies in Life Cycle Impact Assessment (LCIA) generally cover larger scales (100 km to global), are screening oriented (fast evaluation of large numbers of chemicals), and involve multimedia fate and transport and multi-pathway exposure.[3,13,14] Their lower spatial resolution, however, limits their ability to account for important spatial variability in parameters (e.g. population or agricultural production densities), in particular in the vicinity of emission sources.

To better evaluate exposure to chemical emissions, we need both the ability to account for relevant spatial variability (e.g. around emissions sources or highly populated areas), and enough flexibility to easily change the region(s) of focus.

To meet these requirements, we consider three categories of spatial models as previously analyzed and compared.[3,5,13,15,16] The first, *chemical transport models* (CTMs), consists of models with generally high spatial or temporal resolution, such as the atmospheric chemical transport and air quality models AERMOD,[17,18] ADMS,[19] and GEOS-Chem.[20] While being essentially focused on atmospheric fate and transport of chemicals, some CTMs were extended to take other media into account, e.g. CMAQ,[21] ECHAM 5,[22] and EMEP-MSCE.[23] Models from this category are generally slow at runtime (a few minutes up to days per pollutant/scenario), data/parameters intensive, rarely global, and often model other media only as reservoirs, sometimes with no reverse flows (e.g. deposition without re-volatilization) or exchanges between reservoirs. Moreover, none of these models covers sufficiently or at all population multi-pathway exposure.

The second category of models consists of *spatialized multimedia box models*, accounting for intra- and inter-media processes generally described by first order differential equations, such as ChemRange,[24] Globo-POP,[25] ClimoChem,[26] G-CIEMS,[27] NIAES-MMM-Global,[28] BETR variants,[29–33] and IMPACT variants.[34–37] These models were often parameterized using outputs of models from the first category (e.g. meteorology from GEOS-Chem for parameterizing IMPACT). While well suited for screening large numbers of chemicals, these models have other limitations. They are essentially designed with fixed grids and specific parameterizations. This makes it difficult and time consuming to change the region of focus or grids resolution and geometry, which is essential for capturing project-specific relevant spatial variability. Large and especially global scale multimedia box models are generally low

resolution, which introduces artifacts (Section S1.1, SI) such as artificial dilution or lack of discrimination power between sources and receptors, and prevents an accurate characterization of regions of interest. In addition, most of the global spatial multimedia models do not cover multi-pathway human exposure.

A third category of GIS-based models has emerged more recently,[38] which can achieve very high spatial resolution. When based on raster algebra, however, they can become data and computation intensive, especially in contexts where both global scale coverage and high resolution at specific locations are required. While multi-media coupling (with feedback) is possible, it remains rather complicated to achieve, especially if required on a per study basis.

In conclusion, there is a need for a multimedia modeling framework that combines the advantages of the three aforementioned model categories, while overcoming the limitations of decoupled environmental media and fixed grid geometries. Addressing this need could be achieved by nesting or cascading existing models, but this would come at the expense of several person-months of work per specific location and study context. Such framework must therefore be global, account for multimedia fate and transport and multi-pathway exposure, and be flexible enough to enable high resolution at study-specific locations of interest (automatizing GIS processing).

To address this need, we propose a flexible multi-scale, spatially explicit multimedia modeling framework, suited for assessing and comparing human intake from local to global scales for a wide range of emission scenarios and sources. More specifically, we aim to: (1) present the structure of the newly developed *Pangea* modeling framework, (2) evaluate local to global fate and exposure associated with emissions of selected pollutants, analyzing the magnitude of population intakes as a function of chemical properties and distances for 126

116 point sources in France, and (3) simulate emissions from ~10,000 point sources that cover
117 France more densely and build maps of related global intake fractions.
118
119

# METHODS

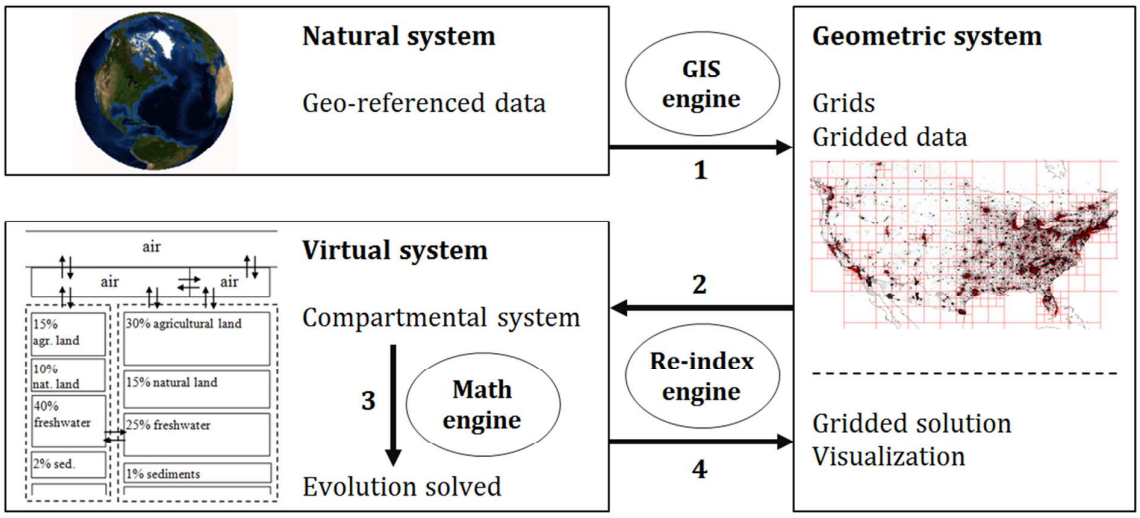
## The *Pangea* Modeling Framework

Conceptually, *Pangea* is a compartmental modeling framework (2nd model category) that incorporates a GIS engine (3rd category), and aims at making a step towards CTMs (1st category) spatial resolutions, while maintaining the ability to perform fast fate and exposure evaluations.

**Model structure** – Multimedia "box" models consist of a set of compartments representing environmental media (e.g. atmosphere, fresh water, specific land covers), associated with a set of environmental processes models (EPMs).[3,13,39–42] EPMs describe interactions between media, i.e. diffusive and advective flows, as well as internal processes such as degradation. *Pangea* extends this structure with the ability to create multi-scale grids and to project spatial data at runtime. This allows building grids adapted to the specifics of each study and, in our context, enables the comparison of local and global impacts of pollutant emissions within the same consistent framework.

The structure of *Pangea* is outlined in Figure 1. A GIS engine based on MATLAB and Python/ArcGIS builds global 3D multi-scale grids, defines geometric and topological parameters, and grids/projects geo-referenced data associated with the *natural system* (e.g. meteorology and terrestrial coverage). This process yields a *geometric system* of grid cells with homogeneous and inhomogeneous content; terrestrial cells, for example, are composed of several land covers and fresh water. This system is transformed into a *virtual system* of homogeneous compartments using a "re-indexing" engine that splits inhomogeneous cells into homogeneous components based on the proportion of medium types. The virtual system is well suited for defining a compartmental system[43] and a set of first order differential equations that describe the evolution of the mass of pollutant in compartments. While *Pangea*

146 can solve the dynamics, we focus this study on the steady-state solution of a linear system
147 with constant coefficients. The solution, the compartmental distribution of environmental
148 concentrations, is combined with exposure data to yield a distribution of human intakes.
149 Ultimately, all relevant quantities (e.g. concentrations and intakes) in the virtual system are
150 re-indexed back to obtain spatial distributions in the initial geometric system.
151




152

153 **Figure 1.** Structure of the *Pangea* modeling framework illustrating the sequence of
154 operations and utilized engines.
155

156 The media considered in this study are air, sea water, sediments, fresh water, natural land, and
157 agricultural land; the latter three resulting from a *Pangea*-specific aggregation of the 22
158 GlobCover types of land cover (Table 1). The geometric system consists of the following
159 grids:

160 - A 3D atmospheric grid of 17 layers covering altitudes up to ~15 km.
161 - A terrestrial grid of clusters delineated by watersheds delivering to a global ocean.
162 - A freshwater sediments grid compatible with the terrestrial grid.

163 An additional grid, called *results grid*, is built for expressing results on a common grid for
164 comparison.
165

**Chemical fate and transport modeling** – Each spatial parameter describing a grid cell property (e.g. volume) or a projected dataset (e.g. population per grid cell) is expressed as a vector whose size equals the number of cells of the associated grid. Similarly, each parameter that describes an interaction between two grids (e.g. advection between atmospheric grid cells) is expressed as a matrix whose size is defined by the size of the grid(s) involved (e.g. $n$ atmospheric cells × $m$ terrestrial cells). These grid-specific vectors and matrices characterize the geometric system. To efficiently parameterize a global mass balance equation (spatializing EPMs developed for working with abstract geometries), we perform a re-indexing operation that splits vectors and matrices from the geometric system according to the proportion of each medium in each grid cell, and transforms them into block vectors/matrices structured by medium type in a common basis. Re-indexed vectors and matrices, whose elements are associated with homogeneous compartments, define the *virtual system*. Its size, noted $n^v$, corresponds to the total number of compartments ($n^v = n^v_{air} + n^v_{fresh\ water} + ..$). We note $n^{es}$ the number of emission scenarios, where each emission scenario defines a distribution of emissions. We restrict the description of the virtual system dynamics to a system of autonomous first order linear differential equations with constant coefficients:

$$\frac{\mathrm{d}\mathbf{M}^v(t)}{\mathrm{d}t} = \mathbf{K}\,\mathbf{M}^v(t) + \mathbf{S}^v \qquad (1)$$

with $\mathbf{S}^v \in \mathrm{P}^{n^v \times n^{es}}$ an array of vectors of constant emission scenarios [kg s$^{-1}$] written in column, $\mathbf{M}^v(t) \in \mathrm{P}^{n^v \times n^{es}}$ the corresponding array of vectors of masses of chemical [kg] at time $t$ [s], and $\mathbf{K} \in \mathrm{P}^{n^v \times n^v}$ a matrix of transfer and elimination rate coefficients [s$^{-1}$]. $\mathbf{K}$ is a sparse matrix, with dimensions typically in the range of 70,000 × 70,000 to 500,000 × 500,000; it can be factorized, which allows to parallelize the solving of Equation 1 in cases where $n^{es}$ is large (e.g. ~10,000 as in the second part of the study). While being mathematically simple and easy to solve numerically, Equation 1 is complex to build, because components of $\mathbf{K}$ depend on spatial parameters (e.g. contact surface areas between mismatching cells

geometries, volumes, projected spatial datasets), which are specific to each configuration of grids. The consequent work (several person-months when done manually) is usually performed only once for models with fixed grids. The GIS and computation engines integrated in *Pangea*, instead, address this challenging task dynamically, at runtime, with a re-projection of all spatial data (e.g. land cover, demographic data, atmospheric flows, flows through the freshwater network) each time a new geometry is generated. A set of EPMs parameterized by projected spatial parameters, environmental non-spatial parameters, and physicochemical parameters, is used to compute the components of the **K** matrix. EPMs are specific to each medium for elimination processes, to each pair of media for transfer processes, and to each exposure pathway. EPMs are currently based on IMPACT 2002[34,35] and USEtox,[44] adapted for taking spatial data into account. A description of media, processes, and EPMs is given in Section S2 (SI).

The solution of Equation 1 is the distribution of masses of chemical as a function of time, expressed in the virtual system, $\mathbf{M}^{v}(t)$. *Pangea* can provide steady-state and dynamic solutions; for the scope of this study, we focus on the steady-state solution, $\mathbf{M}_{ss}^{v}$. This solution of the fate and transport is back re-indexed to relevant grids to get e.g. $\mathbf{M}_{ss}^{atm,L1}$, the spatial distribution of masses in the first layer of the atmospheric grid, or $\mathbf{M}_{ss}^{fw}$, the spatial distribution of masses in the global fresh water network.

**Human exposure modeling** – Exposure pathways considered in this study are inhalation and ingestion of freshwater and food (fish, meat, milk, below-ground produce, and above-ground produce). The population intake through both inhalation and ingestion is defined as:

$$\mathbf{IN}_{ss}^{v} = f(\mathbf{p}^{v}, \mathbf{IR}^{v}, \mathbf{BAF}^{v}, \mathbf{C}_{ss}^{v}) \quad (2)$$

where $\mathbf{C}_{ss}^{v}$ is the array of environmental concentrations corresponding to $\mathbf{M}_{ss}^{v}$, $\mathbf{BAF}^{v}$ is an array of generalized bioaccumulation factors (that can be interpreted as an application from

environmental concentrations to concentrations in substrates: air, water, and food items), $\mathbf{IR}^{v}$ is an array of generalized individual intake rates of substrates, $\mathbf{p}^{v}$ is a vector of population counts, $f$ is an appropriate product operation, and $\mathbf{IN}_{ss}^{v} \in \mathrm{P}^{n^{v} \times n^{es} \times n^{ep}}$ is a 3D array of population intakes [kg $s^{-1}$] with $n^{ep}$ the number of exposure pathways (inhalation of air, ingestion of drinking water, ingestion of different food items). Finally, the population intake fraction (iF) is defined as:

$$(\mathbf{iF}_{ss}^{v})_{ijk} = \left( \sum_{r=1}^{n^{v}} \mathbf{S}_{rj}^{v} \right)^{-1} (\mathbf{IN}_{ss}^{v})_{ijk} \quad (3)$$

Fixing the emission scenario (index $j$, 2nd dimension), this formulation allows to aggregate population intake fractions over compartments (sum over index $i$, 1st dimension) and/or exposure pathways (sum over index $k$, 3rd dimension). *Pangea* considers exposure through inhalation and ingestion of different food items and drinking water. We define the total *population intake* and *intake fraction* as the sum of all intakes and intake fractions related to ingestion. Inhalation and ingestion intake fractions cannot be summed; the former occurs locally, whereas the latter happens wherever production items are consumed ultimately. The missing link would be a trade model similar to the one implemented in IMPACT World.[45] However, currently no food trade model is implemented in *Pangea* because of the difficulty to parameterize it globally at (multi-scale) higher resolution than the country level.

**Input datasets** – Table 1 lists the main data sets involved in the parameterization. The GIS engine performing data projection at runtime makes the selection of data sets flexible; it can be study-specific and evolves with growing data availability.

**Table 1**. Major datasets involved in the process of grids creation, parameters for the EPMs, and for computing population exposure. Section S3.1 (SI) provides a more detailed description of main datasets.

| ***Dataset*** | ***Description*** |
|---|---|

| | |
|---|---|
| GEOS-Chem[20] | Yearly bidirectional averages on a 6 hours basis were computed based on a GEOS-Chem run for the reference year 2005. |
| GlobCover[46] | © ESA 2010 and UCLouvain. The 22 land cover categories are aggregated into *Pangea*-specific categories. |
| WWDRII[47] | World Water Development Report II. Gridded $0.5° \times 0.5°$ water network used as a basis for computing flows between watersheds, as adapted by Helmes et al. (2012).[48] |
| FAOSTAT | Global food production per country in 2009. |
| USEtox[44] | Basis for selected intra- and inter-compartment processes and default parameters. |
| IMPACT 2002[34,35] | Basis for selected intra- and inter-compartments processes and default parameters. |
| CIESIN | Raster of population counts for 2005. |
| ISWA[49] | Waste-to-Energy, State-of-the-Art-Report, 6th Edition, August 2012.<br>Location and capacity of 126 waste-to-energy-plants in France. |
| ESRI | Data and Maps 9.3 for country boundaries. |

239

240

## Local to global scale population exposure

**Case study design** – The comparison of local and global intakes of selected chemicals is based on atmospheric emissions from 126 point sources representing waste-to-energy plants' stacks located across France, mapped using black dots in Figure 2. Source locations were determined based on the ISWA report.[49]

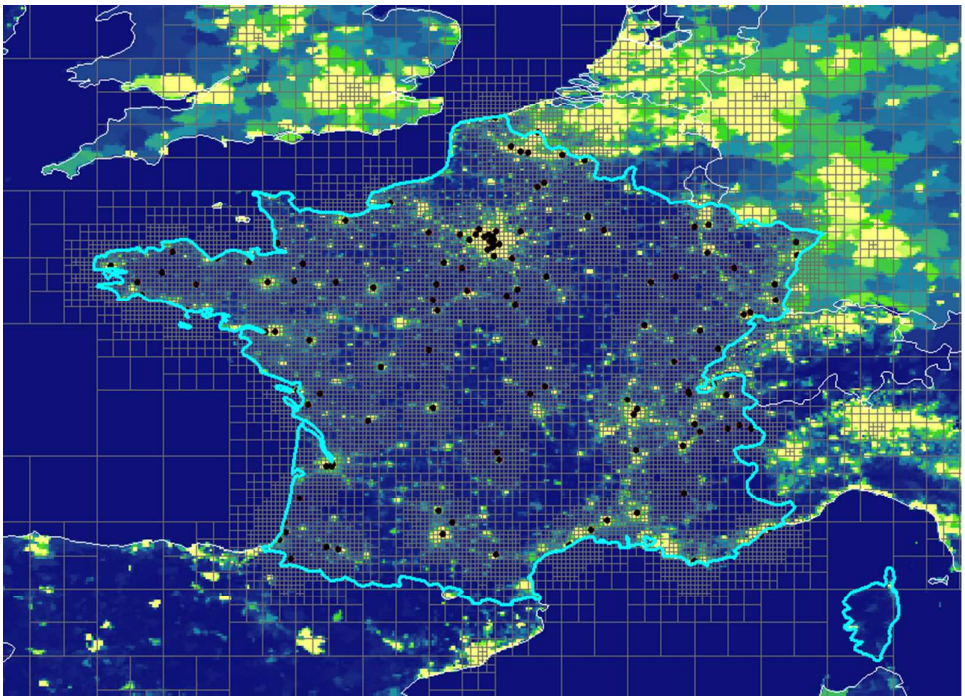


**Figure 2.** 126 point sources, representing waste-to-energy plants' stack, overlaying a map of population density that shows the spatial distribution of the population

Three substances were selected to cover distinct dominant impact pathways, and to represent different classes of pollutants: 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) for dioxins and furans, benzene for volatile organic compounds (VOC), and benzo[*a*]pyrene (B[*a*]P) for polycyclic aromatic hydrocarbons (PAHs). The main properties of these substances are given in Section S3.2 (SI). We first analyze the spatial distribution of concentrations and population intakes for a single "test" atmospheric source located in the North-East of France, illustrating via a series of maps how the model performs and determines intake fractions.

Key to this study is the creation of multi-scale grids with higher resolution over regions of interest, such as the location of emission sources and regions with large populations. The mechanism for building multi-scale grids involves a *refinement potential* (RP): a raster whose pixels quantify the need (potential) for grid refinement. The RP used for this study builds mainly on two components: a raster of inverse distances from sources and a raster of population counts. Starting from a background, low resolution grid, the RP is used as a basis for iteratively refining each cell whose integrated potential is above a defined threshold. The outcome of this refinement method is shown in Figure 2. The grid has a higher resolution around points of emission (approximately 7 km × 7 km cells), which decreases with distance from sources. The resolution is also constrained by the presence of large populations. The first layer of the atmospheric grid is identical to the results grid; the other 16 atmospheric layers have resolutions decreasing with altitude. The terrestrial grid is made of watersheds obtained from clustering the $0.5° \times 0.5°$ gridded water network defined by WWDRII.

**Radial distribution of iFs** – We compute spatial distributions of iFs through inhalation and total ingestion (aggregated over all ingestion pathways) and summarize them radially, for each substance and each of the 126 point sources. Radial summaries are integrated and we present a statistic (curves of percentiles) of the 126 integrals, as well as a statistic normalized by global iFs.

## Maps of global iFs for a large number of emission locations

Maps of global iFs associated with a large number of emission locations in France are obtained by repeating the above study for ~10,000 locations, for each of the three substances. Global/total iFs through all exposure pathways are computed per exposure route, and associated with each emission location.

# RESULTS

## Flexible model parameterization

Using a refinement potential fitting the needs of this study, it takes *Pangea* about 5 minutes on a laptop (Xeon CPU, 32GB RAM) to create a full spatial multimedia model: build grids, connections, and project data. It takes less than 10 seconds per substance to parameterize and run emissions from 126 sources. Exports to data files and maps can then take from a few minutes to hours depending on the level of output desired, mainly because creating maps with thousands of polygons is time consuming.

## Local to global intakes of pollutants

Spatial distributions of environmental concentrations are computed for the three substances. To illustrate the approach underlying the 126 and then ~10,000 scenarios, Figure 3 shows, for each substance, environmental concentrations $[\mathrm{mg/m^3}]$ in air (layer 1), fresh water, and agricultural soils, and the population intake fraction surface densities $[1/\mathrm{m^2}]$ ($[\mathrm{kg}_{taken\ in}/\mathrm{kg}_{emitted}/\mathrm{m^2}]$) through inhalation and total ingestion (i.e. via all ingestion-related pathways), for a unit emission (1 $\mathrm{kg/s}$) at a single emission source located in La Veuve, in the North East of France (indicated by a white cross).

Figure 3.A shows that plumes are stretched in the prevailing wind direction, with atmospheric concentrations decreasing with distance from the source. The size of the plume reflects the substance-specific persistence in air, with long-range transport for benzene (half-life in air of 8.7 days), and a much more localized plume for B[$a$]P which has the shortest half-life in air (0.21 day). The contrast between freshwater concentrations (Figure 3.B) and concentrations in agricultural soils (Figure 3.C) shows that transport through the hydrological network is relevant, especially for less volatile substances. The concentrations in freshwater of Figure

306 3.B reflect the combination of atmospheric transport, deposition, and runoff. The Seine River
307 (medium size with 500 $m^3/s$ average discharge, flowing through Paris), whose watershed is
308 fully under the plume, drives high concentrations West, in a direction opposite to the plume.
309 The Rhine River (large with 2000 $m^3/s$ average discharge, flowing through Germany and
310 The Netherlands), whose watershed is only partly under the plume, has more dilution power
311 and leads to concentrations lower than over its surrounding watershed. Due to its high
312 volatility, benzene concentrations in water and soil are as expected orders of magnitude
313 inferior to TCDD for the same emission scenario.
314
315 Figures 3.D/E show population iFs as a surface density. While the unit $[1/m^2]$ ($[kg_{taken\ in}/$
316 $kg_{emitted}/m^2]$) is more difficult to interpret than a unit-less fraction, it facilitates interpreting
317 maps independently of multi-scale grid cells sizes. While the inhalation intake fraction
318 density of a grid cell refers to the amount of contaminant inhaled by the actual population
319 within that cell, the ingestion intake fraction density refers to the amount of contaminant
320 ingested by the global population via consuming the agricultural produce and fish grown
321 within that cell. Densities of the inhalation or ingestion intake fractions can be integrated over
322 any area to compute (or evaluate visually) the intake faction associated with this area.
323
324 Figure 3.D shows the intake fraction through inhalation; it is driven by the plume and the
325 distribution of population. Regions under the plume, even far from the emission source, in
326 Belgium, The Netherlands, and Germany, represent a large share of the global intake fraction.
327 Global intake fractions through inhalation for benzene, TCDD, and B[a]P are $1.4\times10^{-5}$,
328 $4.9\times10^{-6}$, and $9.5\times10^{-7}$, respectively, which is consistent with their decreasing persistence in
329 air.
330

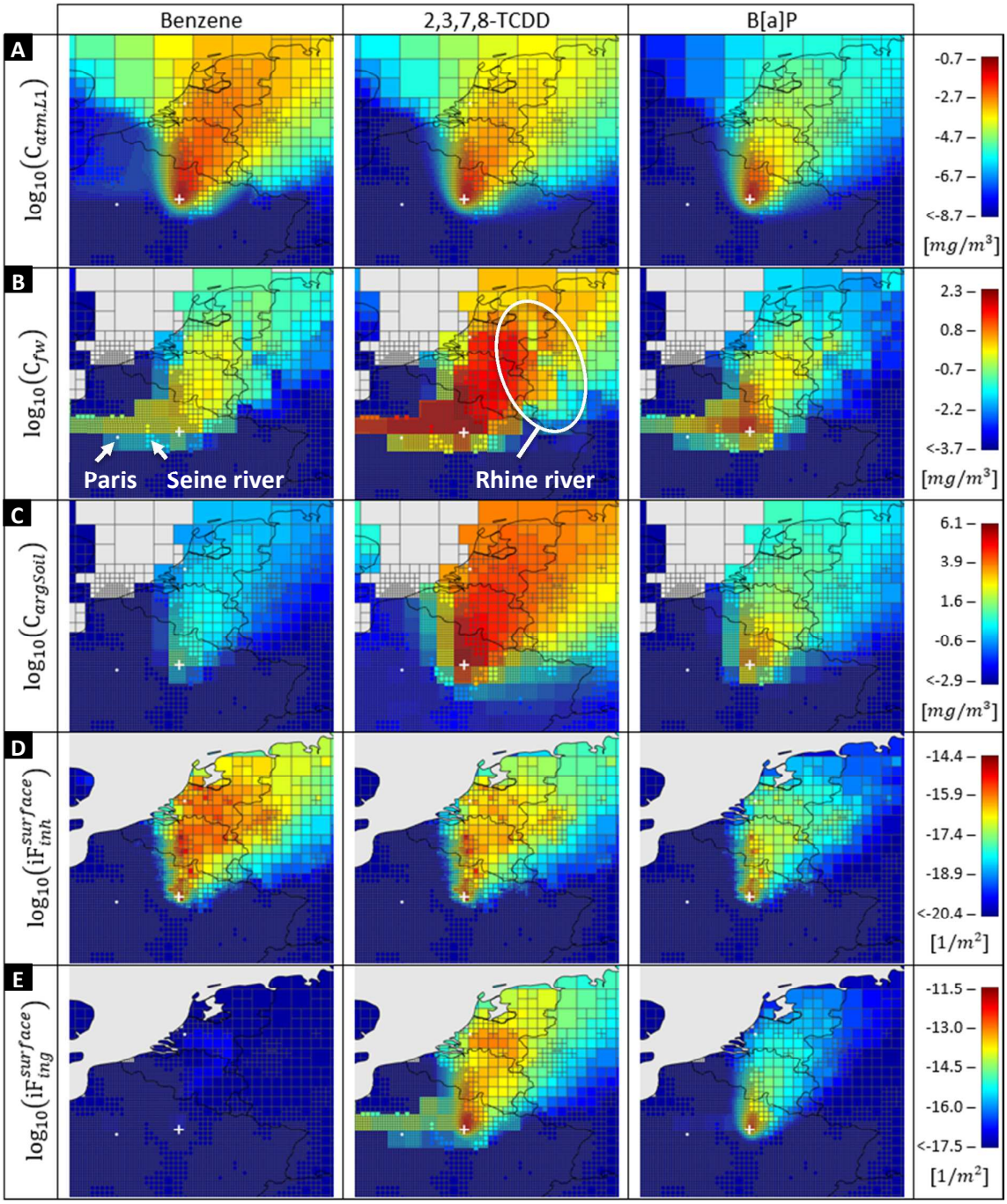
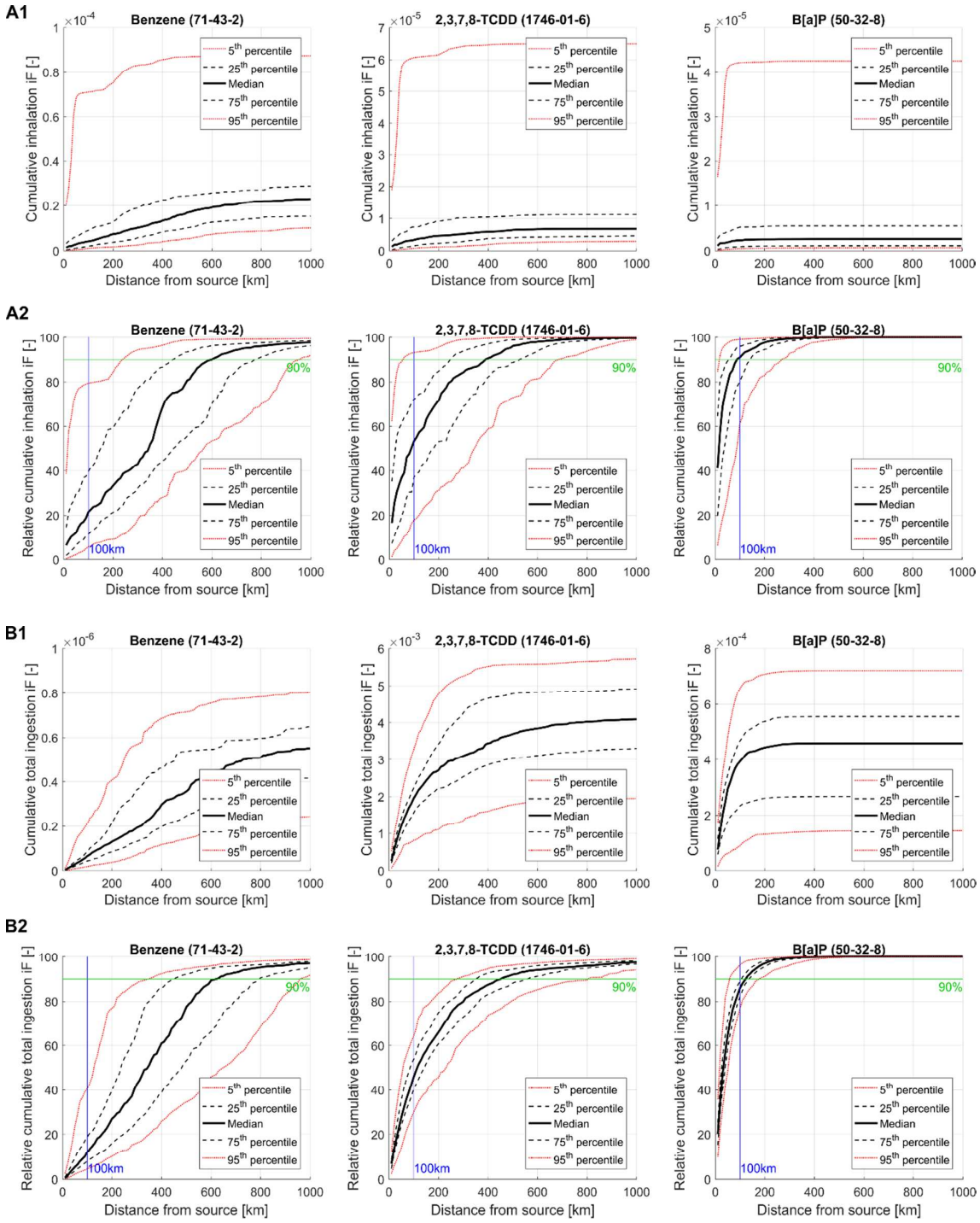



**Figure 3.** Environmental concentrations [mg/m$^3$] at steady-state and subsequent population intake fraction surface densities [1/m$^2$] of selected substances, resulting from a unit (1 kg/s) atmospheric emission at a point source in La Veuve, France (white cross). **A.** Concentration in first layer of atmospheric grid. **B.** Concentration in fresh water. **C.** Concentration in agricultural soils. **D.** Population intake fraction density through inhalation. **E.** Population intake fraction density through total ingestion.

Figure 3.E shows population intake fraction densities through total ingestion. While the global ingestion iF of a volatile substance like benzene ($5.4\times10^{-7}$) is negligible and two orders of magnitude smaller than the global iF though inhalation, global ingestion iFs of TCDD and B[a]F are substantially larger, with $4.9\times10^{-3}$ (main contributions: fish $1.4\times10^{-3}$ and above-ground produces $2.3\times10^{-3}$) for TCDD and $4.8\times10^{-4}$ (main contribution: above-ground produces $4.2\times10^{-4}$) for B[a]F. The relative contribution through the Seine River is more important for TCDD than for B[a]P, which reflects the highest bioconcentration factor in fish and therefore the contribution of fish consumption for TCDD. The ingestion iF density of TCDD gets first reduced as we move away from the source, but increases again over the farther but intensive agriculture producing regions of The Netherlands.

Spatial distributions of the inhalation or ingestion iFs can be summarized radially from any point source and integrated, to express a cumulative population intake as a function of the distance from source. We performed this computation for the 126 emission sources shown in Figure 2, and built a statistic of all radial summaries. Figure 4.A shows the distribution of the cumulative population iF by inhalation, as absolute (A1), and relative (A2, as percent of global iF) values, with 95$^{th}$ percentiles of global inhalation iFs for the three substances between $4.5\times10^{-5}$ to $9\times10^{-5}$. The ratio between the 5$^{th}$ and the 95$^{th}$ percentiles of absolute inhalation iFs amounts to a factor 9 for benzene, and up to a factor 64 for B[a]P that undergoes coarser spatial variations related to the lower persistence in air and more condensed plume around the source. Analyzing the relative iF plot for benzene, the median curve indicates ~20% of global iF over a 100 km radius from point sources, and a median distance for getting 90% of the global iF of ~600 km. The corresponding relative values are ~54% at 100 km and ~400 km for 90% for TCDD, and ~90% at 100km for B[a]P, demonstrating well the differences in travel distances between these substances.

366 Figure 4.B shows the distribution of the evolution of the cumulative ingestion iF, in absolute
367 (B1) and relative (B2, as percent of global ingestion iF) values. Figure 4.B1 shows that
368 ingestion iFs span a wider range, from $9\times10^{-7}$ to $6\times10^{-3}$, than inhalation iFs. The ratio between
369 the 5$^{th}$ and the 95$^{th}$ percentiles of absolute ingestion iFs is less than a factor 6 and therefore
370 lower than the inhalation value. This reflects a more uniform spatial distribution of
371 agricultural production compared to the distribution of population that is highly concentrated
372 in large cities. Analyzing the relative iF plot for benzene, we see that the median curve
373 indicates only ~10% of global ingestion iF originating from food and fish within a 100 km
374 radius from point sources (Figure 4.B2), and that the distance for getting 90% of the global iF
375 is ~600 km. More importantly, because ingestion is the dominant exposure route for these
376 substances, respectively ~46% and ~86% of global ingestion iFs for TCDD and B[a]P are
377 generated by food/fish production within 100 km from point sources, with respectively ~460
378 km and ~120 km needed to capture 90% of their global ingestion iFs.
379
380 A sensitivity study summarized in Section S5 (SI) shows that global intake fractions are stable
381 (in $2.5\times10^{-5}$–$5.0\times10^{-5}$ for benzene inhalation, ~$4.0\times10^{-3}$ for TCDD ingestion, and in $4.0\times10^{-4}$–
382 $5.0\times10^{-4}$ for B[a]P ingestion) with respect to the year of the meteorological field data set, the
383 version of GEOS-Chem (GEOS-4, 2005; GEOS-FP, 2013-2016) and the resolution (2°×2.5°
384 only, and 0.25°×0.3125° over Europe nested in a global 2×2.5°). It also shows that fractions
385 taken in at 100 km are within 21.5%–34.7% for benzene inhalation, 45.9%–73.6% for TCDD
386 ingestion, and 86.0%–97.5% for B[a]P ingestion, which defines three disjoint ranges despite
387 the variability.
388




389
390
391
392

**Figure 4.** Absolute and relative cumulative population intake fractions for three substances, as a function of distance from source. Inhalation absolute (A1) and relative (A2), total ingestion absolute (B1) and relative (B2). The vertical blue line identifies the relative value at 100 km distance from the source (e.g. 20% for benzene/inhalation), the horizontal green line identifies the distance for accounting for 90% of the intake (e.g. 600 km from source for benzene/inhalation).

393
394
395
396
397
398
399

## Maps of global iFs for a large number of emission locations

Figure 5 shows maps of global inhalation iFs (5A) and global ingestion iFs (5B) associated with ~10,000 emission locations distributed over France. Explicitly, the value of each pixel is the value of the global cumulated iF associated with an emission source from this pixel, yielding iF maps by source location.

Comparing Figures 5A and 5B, the dominant exposure route is inhalation for benzene, and total ingestion for TCDD and B[a]P. Analyzing the map of global inhalation iF for benzene, we observe high values in regions down-wind from Paris, Lyon, Geneva, and more generally north of France in Belgium, Germany, and The Netherlands. Maps of global ingestion iF for both TCDD and B[a]P match well the distribution of large crops (maize, wheat, etc.) and animal-based agricultural regions in France.[50]

Figure 5C shows distributions of inhalation and ingestion iFs, as well as the continental rural and urban iF archetypes obtained from USEtox (European region 19). There is a good match between distributions of global inhalation iFs for the three substances and the boundaries defined by USEtox continental and urban regions, considering that the ~10,000 emission locations modeled with *Pangea* include cases of extreme urban exposure (emission precisely in the middle of large population). *Pangea* yields respectively one order of magnitude lower total ingestion of benzene and higher total ingestion of TCDD and B[a]P compared to USEtox.

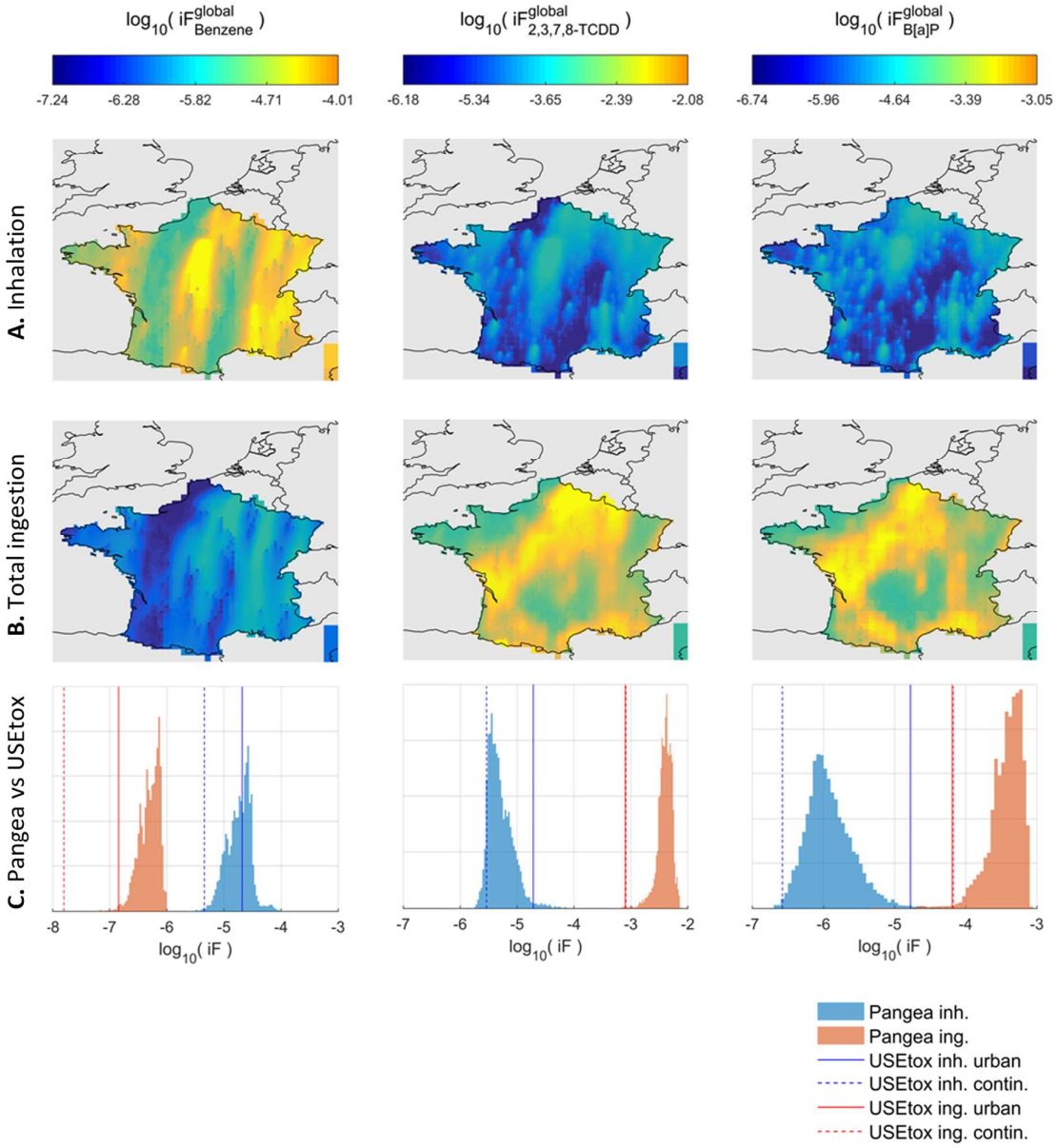



**Figure 5.** Maps of global iF through inhalation and total ingestion associated with ~10,000 emission locations covering France.

# DISCUSSION

We compared local with global population inhalation and ingestion intake fractions for a set of point sources defined by the location of 126 solid waste treatment plants across France. Median curves of cumulative iFs show that for selected volatile and persistent pollutants (benzene in this study), modeling fate and population exposure within a 100 km radius from the source may only account for 20% of the global population intake through inhalation. The 100 km modeling radius also only accounts for 46% of total TCDD ingested, which is again not even half of the total intake.

The systematic sensitivity analysis of iFs for ~10,000 sources covering the entire France demonstrates the feasibility of such studies, and highlights the differences in spatial variations between inhalation that is population driven as compared to ingestion that reflects proximity to agricultural areas. Using both spatial results, Pangea can help classifying regions of emission and identifying worst-case scenarios per substance, or, coupled with a national inventory of emissions, per industrial sector[51].

The two urban and rural continental archetypes of USEtox overlap and reflect well the range of variation in inhalation intake. Further scrutiny is required on the ingestion pathways to further understand and possibly extend the analysis to a receptor perspective by coupling Pangea with multi-regional input-output models relating production locations to food consumption locations.

Two competing factors define the evolution of the population intake radially from the source: decreasing environmental concentrations and increasing exposed populations or agricultural production. Individual intakes decrease in the same fashion as environmental concentrations; locations distant from the source are therefore characterized by a large cumulative exposed

453 population intake of people with small individual intakes. The relevance of this situation
454 relates to the dose-response relationship specific to each pollutant (linear/non-linear) and
455 endpoint (e.g. cancer), i.e. the probability of developing a specific adverse effect expressed as
456 a function of the intake dose for a specific pollutant. Of special interest are carcinogenic
457 pollutants, in particular when considered under the U.S. EPA default assumption of a linear
458 mutagenic mechanism, and situations where doses due to background concentrations are
459 already above threshold. In such cases, a large population exposed to low dose increments
460 may lead to as many or more cases of e.g. cancer as a smaller population exposed to higher
461 doses. This suggests that, by discarding distant population exposure to low dose increments,
462 local and regional assessments might only account for a small fraction of global impacts, in
463 particular for mutagenic substances.
464
465 Several limitations apply to *Pangea*. For ingestion of food crops, Pangea is currently based on
466 the same generic plant model as USEtox[52] and relatively simple coupling with main land
467 cover types (section S3.1.2, SI). It would be interesting to extend the model and include more
468 advanced plant models such as those used for pesticide residue modeling,[53,54] in particular for
469 rice crops that strongly modify multimedia transfers and couple these with crop specific land
470 cover types once they will become available at global level.
471
472 Input datasets and grids have different resolutions; datasets with a resolution lower than the
473 finest multi-scale grid cells are interpolated and datasets with higher resolution are
474 summarized by zonal statistics. It is therefore not possible to define a single resolution for
475 *Pangea*. Simulation durations compatible with pollutant screening can only be achieved
476 computing a steady-state solution of a linear model with constant coefficients. To take full
477 advantage of the multi-scale spatialization of *Pangea* and apply it to e.g. hot spot
478 identification, a wider availability of good quality spatialized emission inventories is

479 essential. There is therefore a need to further develop inventories for a larger number of
480 substances, building e.g. on the work of Breivik et al.[12] for PCBs or Hodges et al.[55] for
481 chemicals in household products. Further, despite the technical ability of the model to build
482 study-specific grids with arbitrarily high resolution, an absolute limit of resolution is defined
483 by the scale at which dynamic and non-linear effects become significant. Finally, steady-state
484 is assumed in the current study. The steady-state assumption is equivalent to a long-term
485 integration of masses and intakes[56] as long as rate coefficients are constant. Such rates can
486 however vary with e.g. temperature and the OH radical concentration determining
487 atmospheric half-lives,[57] whose variability should be accounted for in future research.
488
489 Another need is to perform a systematic model evaluation. A first model evaluation against
490 measured data was performed in the frame of an analysis of freshwater concentration of
491 household and personal care product chemicals in Asia. Wannaz et al.,[58] showing a good
492 agreement for triclosan in freshwater (Pearson r = 0.82, for 253 monitored values covering 12
493 streams), but a significant model underestimation for MeP in sediment that might be due to
494 potential non-considered natural sources. Comparison to measured data were carried out for
495 industrial sources of in Australia, but the lack of an overall inventory only enabled to check
496 that estimated air concentration due to industrial sources were below observed measurements.
497 The condition for an effective evaluation is to have a comprehensive emission inventory from
498 all punctual and diffuse sources over the entire region of influence and corresponding
499 monitored data.
500
501 In conclusion, *Pangea* is an innovative global spatial multimedia fate and multi-pathway
502 exposure modeling framework, able to build study-specific multi-scale grids designed for
503 exploiting high-resolution spatial data sets for characterizing population exposure (intake
504 fraction per pathway) from the vicinity of emission sources to the globe. Based on GIS and

numeric engines and peer-reviewed EPMs, it efficiently transforms geo-spatial, geometric, and topologic data and parameters into a fully coupled compartmental model, and implements methods for simulating large numbers of emission scenarios in parallel. Its flexibility allows focusing on large numbers of point and diffuse sources, allowing for a better characterization of exposure to inform risk- and impact-assessments. It also opens the possibilities to look at exposures not only from an emitter perspective as described this paper, but also from a receptor perspective looking at all sources contributing to exposure in a receptor location.[51] The Pangea model can be accessed via http://www.pangea-model.org.

513 # SUPPORTING INFORMATION

514 Descriptions of environmental media, datasets, and main chemical properties along with
515 details of the computation of human exposure are available as Supporting Information for this
516 article. This material is available free of charge via the internet at http://pubs.acs.org.

517

518 # ACKNOWLEDGMENTS

519 We thank Dr. Shanna Shaked for her scientific support during model development. This work
520 was in part supported by Veolia Environment and by the Marie Curie project Quan-Tox (grant
521 agreement no. 631910) funded by the European Commission under the Seventh Framework
522 Programme.

523

524

# REFERENCES


(1) Zou, B. How should environmental exposure risk be assessed? A comparison of four methods for exposure assessment of air pollutions. *Environ. Monit. Assess.* **2010**, *166* (1–4), 159–167.

(2) Silverman, K. C.; Tell, J. G.; Sargent, E. V; Qiu, Z. Comparison of the Industrial Source Complex and AERMOD dispersion models: Case study for human health risk assessment. *J. Air Waste Manag. Assoc.* **2007**, *57* (12), 1439–1446.

(3) EEA. Feasibility study: modelling environmental concentrations of chemicals from emission data. *Eur. Environ. Agency* **2007**.

(4) Lohman, K.; Seigneur, C. Atmospheric fate and transport of dioxins: Local impacts. *Chemosphere* **2001**, *45* (2), 161–171.

(5) Hollander, A.; Scheringer, M.; Shatalov, V.; Mantseva, E.; Sweetman, A.; Roemer, M.; Baart, A.; Suzuki, N.; Wegmann, F.; Van De Meent, D. Estimating overall persistence and long-range transport potential of persistent organic pollutants: A comparison of seven multimedia mass balance models and atmospheric transport models. *J. Environ. Monit.* **2008**, *10* (10), 1139–1147.

(6) Scheringer, M.; Jones, K. C.; Matthies, M.; Simonich, S.; Van De Meent, D. Multimedia partitioning, overall persistence, and long-range transport potential in the context of pops and pbt chemical assessments. *Integr. Environ. Assess. Manag.* **2009**, *5* (4), 557–576.

(7) Zarfl, C.; Scheringer, M.; Matthies, M. Screening criteria for long-range transport potential of organic substances in water. *Environ. Sci. Technol.* **2011**, *45* (23), 10075–10081.

(8) Scheringer, M.; Strempel, S.; Hukari, S.; Ng, C. A.; Blepp, M.; Hungerbuhler, K. How many persistent organic pollutants should we expect? *Atmos. Pollut. Res.* **2012**, *3* (4), 383–391.

(9) Kuramochi, H.; Takigami, H.; Scheringer, M.; Sakai, S.-I. Estimation of physicochemical properties of 52 non-PBDE brominated flame retardants and evaluation of their overall persistence and long-range transport potential. *Sci. Total Environ.* **2014**, *491–492*, 108–117.

(10) Mackay, D.; Cowan-Ellsberry, C. E.; Powell, D. E.; Woodburn, K. B.; Xu, S.; Kozerski, G. E.; Kim, J. Decamethylcyclopentasiloxane (D5) environmental sources, fate, transport, and routes of exposure. *Environ. Toxicol. Chem.* **2015**, *34* (12), 2689–2702.

(11) Hageman, K. J.; Bogdal, C.; Scheringer, M. Long-Range and Regional Atmospheric Transport of POPs and Implications for Global Cycling. *Comprehensive Analytical Chemistry*. Elsevier: Department of Chemistry, University of Otago, Dunedin, New Zealand 2015, pp 363–387.

(12) Breivik, K.; Armitage, J. M.; Wania, F.; Sweetman, A. J.; Jones, K. C. Tracking the Global Distribution of Persistent Organic Pollutants Accounting for E-Waste Exports to Developing Regions. *Environ. Sci. Technol.* **2016**, *50* (2), 798–805.

(13) Scheringer, M.; Wania, F. Multimedia models of global transport and fate of persistent organic pollutants. In *The Handbook of Environmental Chemistry*; Springer Press, Ed.; 2003; pp 237–269.

(14) McKone, T. E.; MacLeod, M. Tracking multiple pathways of human exposure to persistent multimedia pollutants: Regional, continental, and global-scale models. *Annual Review of Environment and Resources*. University of California, School of Public Health, Lawrence Berkeley National Lab., One Cyclotron Road, 90R-3058, Berkeley, CA 94720, United States 2003, pp 463–492.

(15) Lammel, G.; Klöpffer, W.; Semeena, V. S.; Schmidt, E.; Leip, A. Multicompartmental

fate of persistent substances: Comparison of predictions from multi-media box models and a multicompartment chemistry-atmospheric transport model. *Environ. Sci. Pollut. Res.* **2007**, *14* (3), 153–165.

(16) Hansen, K. M.; Prevedouros, K.; Sweetman, A. J.; Jones, K. C.; Christensen, J. H. A process-oriented inter-comparison of a box model and an atmospheric chemistry transport model: Insights into model structure using α-HCH as the modelled substance. *Atmos. Environ.* **2006**, *40* (12), 2089–2104.

(17) Cimorelli, A. J.; Perry, S. G.; Venkatram, A.; Weil, J. C.; Paine, R. J.; Wilson, R. B.; Lee, R. F.; Peters, W. D.; Brode, R. W. AERMOD: A dispersion model for industrial source applications. Part I: General model formulation and boundary layer characterization. *J. Appl. Meteorol.* **2005**, *44* (5), 682–693.

(18) Perry, S. G.; Cimorelli, A. J.; Paine, R. J.; Brode, R. W.; Weil, J. C.; Venkatram, A.; Wilson, R. B.; Lee, R. F.; Peters, W. D. AERMOD: A Dispersion model for industrial source applications. Part II: Model performance against 17 field study databases. *J. Appl. Meteorol.* **2005**, *44* (5), 694–708.

(19) Cambridge Environmental Research Consultants. ADMS5 http://www.cerc.co.uk/environmental-software/ADMS-model.html.

(20) Bey, I.; Jacob, D. J.; Yantosca, R. M.; Logan, J. A.; Field, B. D.; Fiore, A. M.; Li, Q.; Liu, H. Y.; Mickley, L. J.; Schultz, M. G. Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation. *J. Geophys. Res. Atmos.* **2001**, *106* (D19), 23073–23095.

(21) CMAS. Community Multiscale Air Quality Model https://www.cmascenter.org/cmaq/.

(22) Stevens, B.; Giorgetta, M.; Esch, M.; Mauritsen, T.; Crueger, T.; Rast, S.; Salzmann, M.; Schmidt, H.; Bader, J.; Block, K.; et al. Atmospheric component of the MPI-M earth system model: ECHAM6. *J. Adv. Model. Earth Syst.* **2013**, *5* (2), 146–172.

(23) EMEP-MSCE http://www.msceast.org.

(24) Scheringer, M. Persistence and spatial range as endpoints of an exposure-based assessment of organic chemicals. *Environ. Sci. Technol.* **1996**, *30* (5), 1652–1659.

(25) Wania, F.; Mackay, D. *The global distribution model. A non-steady state multi-compartmental mass balance model of the fate of persistent organic pollutants in the global environment.*; 2000.

(26) Scheringer, M.; Wegmann, F.; Fenner, K.; Hungerbühler, K. Investigation of the cold condensation of persistent organic pollutants with a global multimedia fate model. *Environ. Sci. Technol.* **2000**, *34* (9), 1842–1850.

(27) Suzuki, N.; Murasawa, K.; Sakurai, T.; Nansai, K.; Matsuhashi, K.; Moriguchi, Y.; Tanabe, K.; Nakasugi, O.; Morita, M. Geo-referenced multimedia environmental fate model (G-CIEMS): Model formulation and comparison to the generic model and monitoring approaches. *Environ. Sci. Technol.* **2004**, *38* (21), 5682–5693.

(28) Wei, Y.; Nishimori, M.; Kobara, Y.; Akiyama, T. Development of global scale multimedia contaminant fate model: Incorporating paddy field compartment. *Sci. Total Environ.* **2008**, *406* (1–2), 219–226.

(29) MacLeod, M.; Woodfine, D. G.; Mackay, D.; McKone, T.; Bennett, D.; Maddalena, R. BETR North America: a regionally segmented multimedia contaminant fate model for North America. *Environ. Sci. Pollut. Res. Int.* **2001**, *8* (3), 156–163.

(30) Woodfine, D. G.; MacLeod, M.; Mackay, D.; Brimacombe, J. R. Development of continental scale multimedia contaminant fate models: Integrating GIS. *Environ. Sci. Pollut. Res.* **2001**, *8* (3), 164–172.

(31) MacLeod, M.; Riley, W. J.; McKone, T. E. Assessing the influence of climate variability on atmospheric concentrations of polychlorinated biphenyls using a global-scale mass balance model (BETR-Global). *Environ. Sci. Technol.* **2005**, *39* (17), 6749–6756.

626 (32) Prevedouros, K.; MacLeod, M.; Jones, K. C.; Sweetman, A. J. Modelling the fate of
627 persistent organic pollutants in Europe: Parameterisation of a gridded distribution
628 model. *Environ. Pollut.* **2004**, *128* (1–2), 251–261.
629 (33) Prevedouros, K.; Jones, K. C.; Sweetman, A. J. European-scale modeling of
630 concentrations and distribution of polybrominated diphenyl ethers in the
631 pentabromodiphenyl ether product. *Environ. Sci. Technol.* **2004**, *38* (22), 5993–6001.
632 (34) Margni, M.; Pennington, D. W.; Amman, C.; Jolliet, O. Evaluating
633 multimedia/multipathway model intake fraction estimates using POP emission and
634 monitoring data. *Environ. Pollut.* **2004**, *128* (1–2), 263–277.
635 (35) Pennington, D. W.; Margni, M.; Ammann, C.; Jolliet, O. Multimedia fate and human
636 intake modeling: Spatial versus nonspatial insights for chemical emissions in Western
637 Europe. *Environ. Sci. Technol.* **2005**, *39* (4), 1119–1128.
638 (36) Humbert, S.; Manneh, R.; Shaked, S.; Wannaz, C.; Horvath, A.; Deschênes, L.; Jolliet,
639 O.; Margni, M. Assessing regional intake fractions in North America. *Sci. Total*
640 *Environ.* **2009**, *407* (17), 4812–4820.
641 (37) Jolliet, O.; Shaked, S.; Friot, D.; Humbert, S.; Schwarzer, S.; Margni, M.
642 Multicontinental long range intake fraction of POPS:importance of food exposure and
643 food exports. *Organohalogen Compd.* **2008**, *70*, 1939–1941.
644 (38) Pistocchi, A. GIS Based Chemical Fate Modeling: Principles and Applications. John
645 Wiley and Sons: Hoboken, New Jersey 2014.
646 (39) Rosenbaum, R. K.; Margni, M.; Jolliet, O. A flexible matrix algebra framework for the
647 multimedia multipathway modeling of emission to impacts. *Environ. Int.* **2007**, *33* (5),
648 624–634.
649 (40) Bachmann, T. M. *Hazardous Substances and Human Health: Exposure, Impact and*
650 *External Cost Assessment at the European Scale*; Elsevier, 2006; Vol. Volume 8.
651 (41) Mackay, D.; Webster, E.; Cousins, I.; Cahill, T.; Foster, K.; Gouin, T. *An Introduction*
652 *to Multmedia Models*; Petersborough, Canada, 2001.
653 (42) MacLeod, M.; Scheringer, M.; McKone, T. E.; Hungerbuhler, K. The state of
654 multimedia mass-balance modeling in environmental science and decision-making.
655 *Environ. Sci. Technol.* **2010**, *44* (22), 8360–8364.
656 (43) Jacquez, J. A.; Simon, C. P. Qualitative theory of compartmental systems. *SIAM Rev.*
657 **1993**, *35* (1), 43–79.
658 (44) Rosenbaum, R. K.; Bachmann, T. M.; Gold, L. S.; Huijbregts, M. A. J.; Jolliet, O.;
659 Juraske, R.; Koehler, A.; Larsen, H. F.; MacLeod, M.; Margni, M.; et al. USEtox - The
660 UNEP-SETAC toxicity model: Recommended characterisation factors for human
661 toxicity and freshwater ecotoxicity in life cycle impact assessment. *Int. J. Life Cycle*
662 *Assess.* **2008**, *13* (7), 532–546.
663 (45) Shaked, S. Multi-Continental Multimedia Model of Pollutant Intake and Application to
664 Impacts of Global Emissions and Globally Traded Goods, The University of Michigan,
665 Ann Arbor, 2011.
666 (46) ESA; GlobCover. GlobCover. ESA, Université Catholique de Louvain 2009.
667 (47) Vörösmarty, C. J. Global water assessment and potential contributions from Earth
668 Systems Science. *Aquat. Sci.* **2002**, *64* (4), 328–351.
669 (48) Helmes, R. J. K.; Huijbregts, M. A. J.; Henderson, A. D.; Jolliet, O. Spatially explicit
670 fate factors of phosphorous emissions to freshwater at the global scale. *Int. J. Life*
671 *Cycle Assess.* **2012**, *17* (5), 646–654.
672 (49) ISWA. *Waste-to-Energy, State-of-the-Art-Report*; 2012.
673 (50) Wikipedia. Agriculture en France.
674 (51) Wannaz, C.; Fantke, P.; Lane, J.; Jolliet, O. Source-to-exposure assessment with the
675 Pangea multi-scale framework - Case study in Australia. *Environ. Sci. Process.*
676 *Impacts* **2017**. doi: 10.1039/C7EM00523G

(52) Rosenbaum, R. K.; Huijbregts, M. A. J.; Henderson, A. D.; Margni, M.; McKone, T. E.; Van De Meent, D.; Hauschild, M. Z.; Shaked, S.; Li, D. S.; Gold, L. S.; et al. USEtox human exposure and toxicity factors for comparative assessment of toxic emissions in life cycle analysis: Sensitivity to key chemical properties. *Int. J. Life Cycle Assess.* **2011**, *16* (8), 710–727.

(53) Fantke, P.; Wieland, P.; Wannaz, C.; Friedrich, R.; Jolliet, O. Dynamics of pesticide uptake into plants: From system functioning to parsimonious modeling. *Environ. Model. Softw.* **2013**, *40*, 316–324.

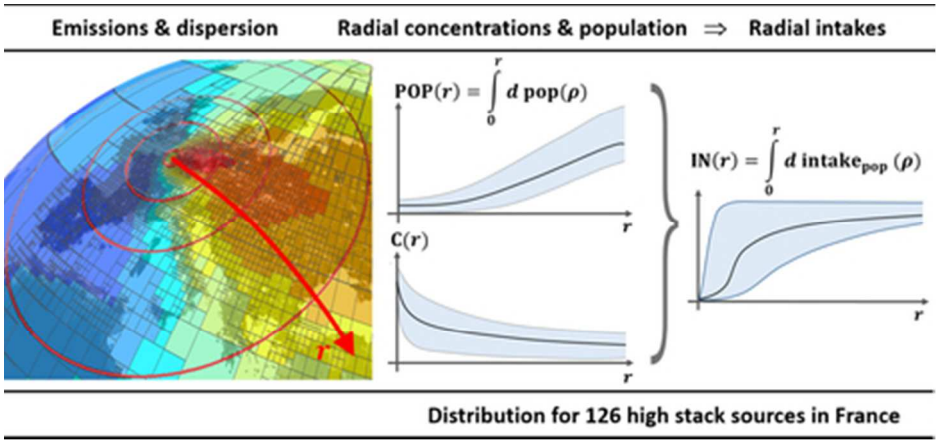
(54) Fantke, P.; Jolliet, O. Life cycle human health impacts of 875 pesticides. *Int. J. Life Cycle Assess.* **2016**, *21* (5), 722–733.

(55) Hodges, J. E. N.; Holmes, C. M.; Vamshi, R.; Mao, D.; Price, O. R. Estimating chemical emissions from home and personal care products in China. *Environ. Pollut.* **2012**, *165*, 199–207.

(56) Heijungs, R. Harmonization of methods for impact assessment. *Environ. Sci. Pollut. Res.* **1995**, *2* (4), 217–224.

(57) Bahm, K.; Khalil, M. A. K. A new model of tropospheric hydroxyl radical concentrations. *Chemosphere* **2004**, *54* (2), 143–166.

(58) Wannaz, C.; Franco, A.; Kilgallon, J.; Hodges, J.; Jolliet, O. A global framework to model spatial ecosystems exposure to home and personal care chemicals in Asia. *Sci. Total Environ.* **2018**, *622*, 410–420.




TOC Art

39x18mm (300 x 300 DPI)